

## RESEARCH LETTER

10.1002/2017GL075933

## Key Points:

- Line by line model coupled to highly accurate multistream solver determines expected instantaneous aerosol radiative effect globally
- Instantaneous aerosol radiative effect and radiative parameterization error evaluated for the native properties and grids of two GCMs
- Error varies spatially and with aerosol optical properties. Global average error of 20% in aerosol absorption seen for one model

## Supporting Information:

- Supporting Information S1

## Correspondence to:

A. L. Jones,  
alj2@princeton.edu

## Citation:

Jones, A. L., Feldman, D. R., Freidenreich, S., Paynter, D., Ramaswamy, V., Collins, W. D., & Pincus, R. (2017). A new paradigm for diagnosing contributions to model aerosol forcing error. *Geophysical Research: Letters*, 44, 12,004–12,012. <https://doi.org/10.1002/2017GL075933>

Received 5 OCT 2017

Accepted 14 NOV 2017

Accepted article online 20 NOV 2017

Published online 7 DEC 2017

©2017. The Authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

## A New Paradigm for Diagnosing Contributions to Model Aerosol Forcing Error

A. L. Jones<sup>1,2</sup>, D. R. Feldman<sup>3</sup>, S. Freidenreich<sup>1</sup>, D. Paynter<sup>1</sup>, V. Ramaswamy<sup>1</sup>, W. D. Collins<sup>3,4</sup>, and R. Pincus<sup>5</sup>

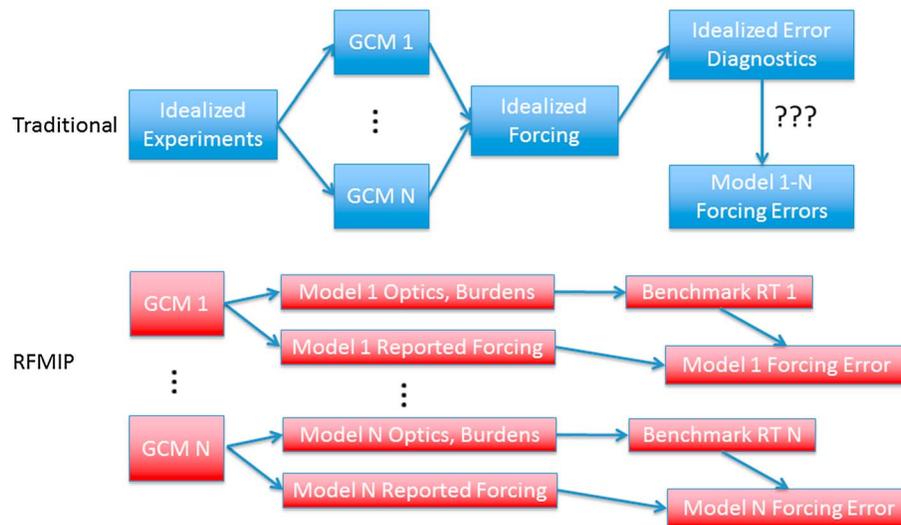
<sup>1</sup>NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA, <sup>2</sup>Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA, <sup>3</sup>Lawrence Berkeley National Laboratory, Berkeley, CA, USA, <sup>4</sup>Department of Earth and Planetary Science, University of California, Berkeley, CA, USA, <sup>5</sup>Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA

**Abstract** A new paradigm in benchmark absorption-scattering radiative transfer is presented that enables both the globally averaged and spatially resolved testing of climate model radiation parameterizations in order to uncover persistent sources of biases in the aerosol instantaneous radiative effect (IRE). A proof of concept is demonstrated with the Geophysical Fluid Dynamics Laboratory AM4 and Community Earth System Model 1.2.2 climate models. Instead of prescribing atmospheric conditions and aerosols, as in prior intercomparisons, native snapshots of the atmospheric state and aerosol optical properties from the participating models are used as inputs to an accurate radiation solver to uncover model-relevant biases. These diagnostic results show that the models' aerosol IRE bias is of the same magnitude as the persistent range cited ( $\sim 1 \text{ W/m}^2$ ) and also varies spatially and with intrinsic aerosol optical properties. The findings underscore the significance of native model error analysis and its dispositive ability to diagnose global biases, confirming its fundamental value for the Radiative Forcing Model Intercomparison Project.

### 1. Introduction

Aerosol radiative forcing and its uncertainty have been implicated as major drivers in the persistently large range in radiative forcing that is reported by climate models (Andreae et al., 2005; Ramanathan & Carmichael, 2008) and in the challenges that climate models have faced in reproducing observed temperature records (Boucher et al., 2013; Santer et al., 2014; Solomon et al., 2011; Stevens, 2015). The climate models reporting results for the Fourth and Fifth Assessment Reports of the Intergovernmental Panel on Climate Change exhibited a direct forcing by natural and anthropogenic aerosols that ranged from  $-0.90$  to  $+0.10 \text{ W/m}^2$  (Forster et al., 2007) and from  $-0.85$  to  $+0.15 \text{ W/m}^2$  (Myhre et al., 2013), respectively. This persistently large range of aerosol radiative forcing, and lingering questions about its sign between Coupled Model Intercomparison Project—Phase 3 (Meehl et al., 2007) and Phase 5 (Taylor et al., 2012) models (CMIP3 and CMIP5, respectively), imply that diagnostics of climate sensitivity over the 1850–2005 historical period are hampered by an imprecise understanding of forcing agents (Stouffer et al., 2017). The  $\sim 1^\circ\text{C}$  of warming seen over that period might be due to large negative aerosol forcing in a high-sensitivity climate or small aerosol forcing in a low-sensitivity climate.

Given the spatial, temporal, and optical heterogeneity of atmospheric aerosols, their actual contribution to forcing is particularly challenging to quantify. Therefore, a systematic evaluation of the impact on simulated climate across the diversity of aerosol distributions resulting from the same experiment is necessary to make substantial progress in resolving this persistent ambiguity in climate sensitivity. To do this, we will evaluate aerosol instantaneous radiative effect (IRE) within models, which is defined as the change in direct, unadjusted radiative flux that would result from instantaneously removing all aerosol (i.e., both natural and anthropogenic) from the atmosphere. IRE differs from direct aerosol instantaneous radiative forcing (IRF), which is normally defined as the impact of only anthropogenic aerosol since 1850 (Myhre et al., 2013). Numerous efforts have been undertaken to examine causes of the spread in IRE or IRF by prescribing aerosol perturbations to each participant model (Mann et al., 2014; Randles et al., 2013; Shindell et al., 2013; Stier et al., 2013). As part of AeroCom, prescribed aerosols were applied to explore the role of radiative transfer (Randles et al., 2013), model component uncertainties (Stier et al., 2013), and microphysics (Mann et al., 2014) in determining IRE error. Additional multimodel exercises with prescribed emissions or climatologies are planned with AerChemMIP (Collins et al., 2017) and the Max Planck Institute Aerosol Climatology (Kinne et al., 2013; Stevens et al., 2017).

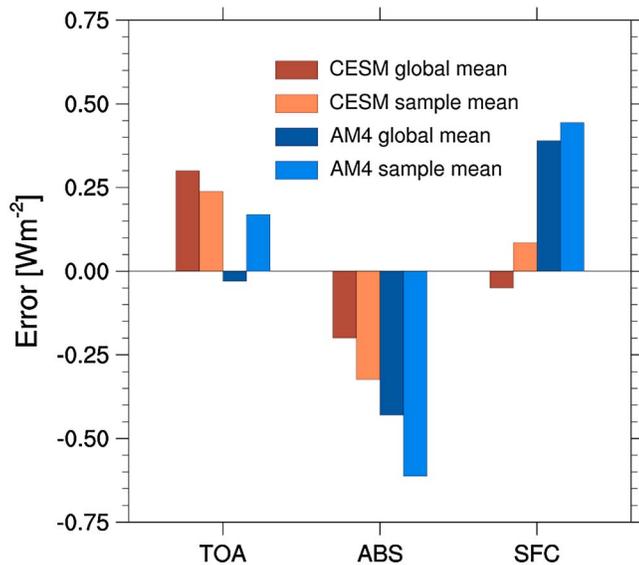


**Figure 1.** Flow charts comparing approach presented here to traditional approaches. The traditional approach uses a set of idealized experiments that yield a corresponding set of idealized errors and associated diagnostics with an unclear relationship with actual model forcing errors. The presented approach, by which models are tested using standard experiments and run natively, reveals the model's native forcing errors.

However, these studies show a large multimodel range in aerosol IRE, and it is difficult to know what portion of this range in aerosol IRE is due to uncertainty in model-specific aerosol, meteorology, and surface albedo properties and their distributions, and what portion is due to error in the radiative parameterizations. While these studies have identified potential sources of model error, bridging the gap to error reduction in the model remains a challenge. In that spirit, we have developed a new paradigm for diagnosing model error due to radiative parameterization that relies on neither a handful of often-idealized profiles nor a climatology prescribed for all participating models. Rather, we request and evaluate detailed diagnostics from participating models as they are run in support of regular contributions to the CMIP archive. We do this to understand model error in aerosol IRE for aerosol conditions native to the model. The schematic in Figure 1 shows this process and how it compares to the previous approaches for determining aerosol IRF.

The large computational expense associated with running a benchmark code on more than just a few sample columns has generally precluded a global benchmark evaluation of general circulation model (GCM) results. However, recent increases in computational power make it possible to run these benchmarking codes globally on a GCM's native grid with that GCM's predicted distributions of meteorology and aerosols. This enables a new, rigorous, and comprehensive approach to quantifying the radiative transfers errors associated with the wide variety of aerosols and climate conditions in that GCM. The goal of this work is to characterize the radiative parameterization error on aerosol IRE for each participant model; this will provide modeling centers with the detailed information necessary to uncover sources of error.

This goal is accomplished by running a line-by-line model coupled to an absorbing and scattering multi-stream radiative transfer solver at every grid point for each GCM on select preindustrial and present dates. The individual model benchmark calculations when compared to that model's native calculations will elucidate that model's radiative parameterization error, which can then be used by the modeling center to determine the relationship between error and the representation of aerosols in the model. Here we will demonstrate preliminary results from this new paradigm of radiative parameterization error quantification of aerosol IRE, especially relative to other approaches, for two sample models: the AM4 model (Donner et al., 2011; Zhao et al., 2016) from the National Oceanic and Atmospheric Administration's Geophysical Fluid Dynamics Laboratory and the National Center for Atmospheric Research's Community Earth System Model (CESM) 1.2.2 (Hurrell et al., 2013). The novel diagnostics produced, coupled with the expertise of each modeling center, will help narrow the persistently large range in aerosol forcing across the multimodel ensemble.



**Figure 2.** Aerosol IRE error metrics at top of atmosphere, within atmosphere absorption, and at the surface for AM4 (blue) and CESM (orange). Errors averaged over four sample columns (light colors) are shown relative to true global mean global errors (dark colors).

## 2. Materials and Methods: Benchmark Setup

The atmospheric, aerosol optical, and boundary conditions involved in internal computation of spectral fluxes as well as the internally computed cloud-free solar fluxes with and without aerosols are produced by each GCM for the spring equinox to ensure global solar illumination, with both models driven by “present-day” aerosol emissions from CMIP5. The distributions of gas concentrations, temperatures, pressures, and geometric thicknesses are input into a line-by-line (LBL) radiative transfer model, the Reference Forward Model (RFM) (Dudhia et al., 2017) to compute gaseous optical depths. RFM queries the HITRAN 2012 spectral database (Rothman et al., 2013) to acquire line strengths at high resolution then reports optical depths at  $1 \text{ cm}^{-1}$  resolution over a spectral range of 1 to  $50,000 \text{ cm}^{-1}$ . This reporting resolution has less than 0.1% impact upon the aerosol IRE compared to  $10^{-4} \text{ cm}^{-1}$  resolution. Those gaseous optical depths, along with aerosol optical depth (AOD), single-scattering albedo (SSA), asymmetry parameter, and boundary conditions from the GCM, are then input to a benchmark quality multistream, doubling and adding solver (Freidenreich & Ramaswamy, 2005; Ramaswamy & Freidenreich, 1991). The solver was run with 16 streams since results converged to within  $0.001 \text{ W/m}^2$  of the 32 stream solution. Radiative transfer calculations are performed at every grid point on instantaneous output. Since grid

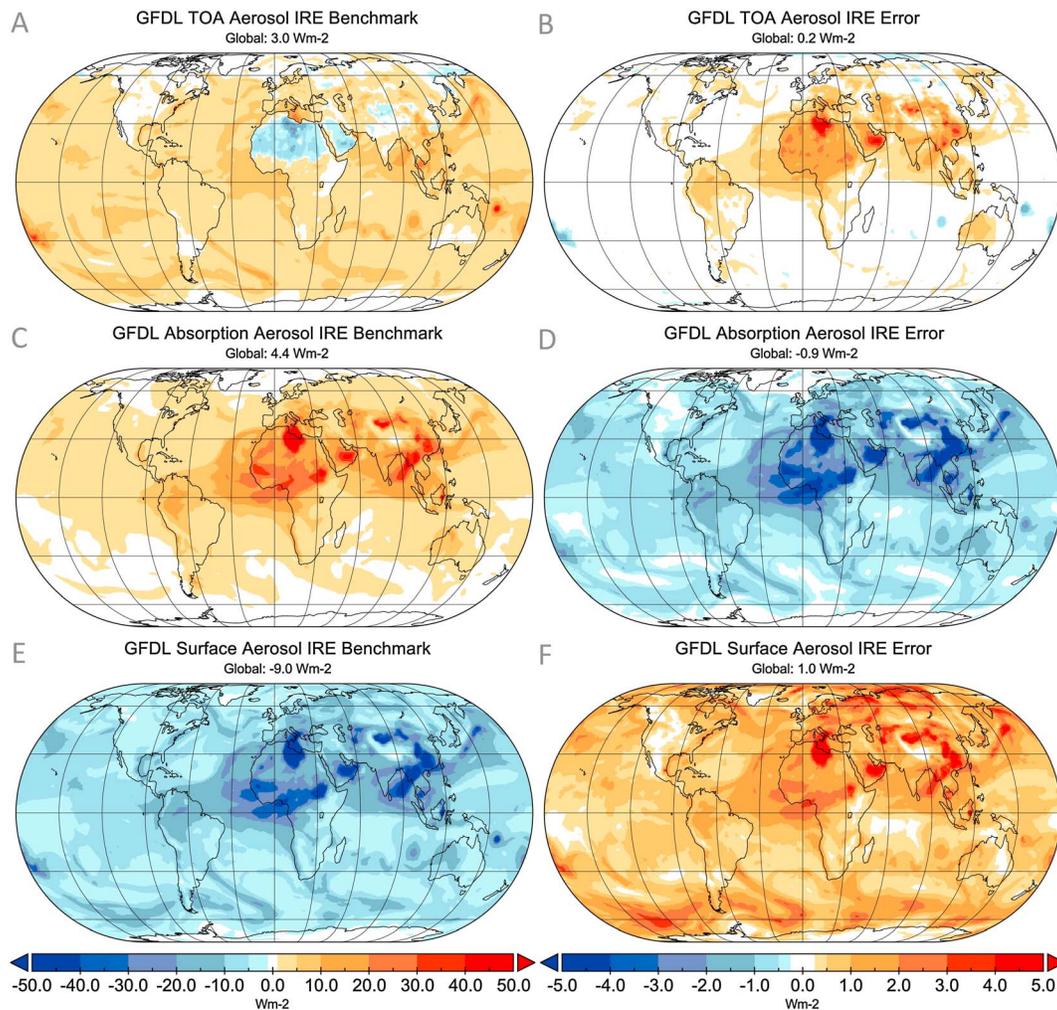
point radiative calculations are independent, the global LBL calculations can easily be parallelized. To calculate global LBL fluxes for a single time step requires about 8,000 computational hours. This method will be used for the Radiative Forcing Model Intercomparison Project-Aerosol Instantaneous Radiative Forcing Component (RFMIP Aerosol-IRF) (Pincus et al., 2016), an endorsed CMIP6 model intercomparison project (Eyring et al., 2016).

The fluxes for all time steps are averaged together to create daily average fields both with and without aerosols. To calculate aerosol IRE the fluxes without aerosols are subtracted from the fluxes with aerosols for three quantities: the net upward flux at the top of atmosphere (TOA), the flux absorbed in the atmosphere, and the downward flux at the surface. To calculate the error in aerosol IRE, we subtract each GCM’s benchmark solution from its internal radiative parameterization calculation of aerosol IRE. Both parameterizations tested here use a two-stream Delta-Eddington approximation and have 10–20 spectral bands for solar radiation. The result represents the error due to the GCM radiative transfer parameterization, since inputs to the benchmark solution are identical to inputs to the native radiative calculations.

This setup has been benchmarked and validated against the Continual Intercomparison of Radiation Codes (CIRC) (Oreopoulos et al., 2012). Specifically, we compared the results of RFM against the Line-by-Line Radiative Transfer Model (Clough et al., 2005; Moncet & Clough, 1997) under cloud-free and cloud-free, aerosol-free conditions. We found differences in top of atmosphere, upwelling diffuse, and surface downwelling direct and diffuse radiation are less than  $1 \text{ W/m}^2$  across the entire spectrum which can be largely attributed to updates in spectroscopy (e.g., Rothman et al., 2013) that have occurred since the CIRC benchmark results.

## 3. Results

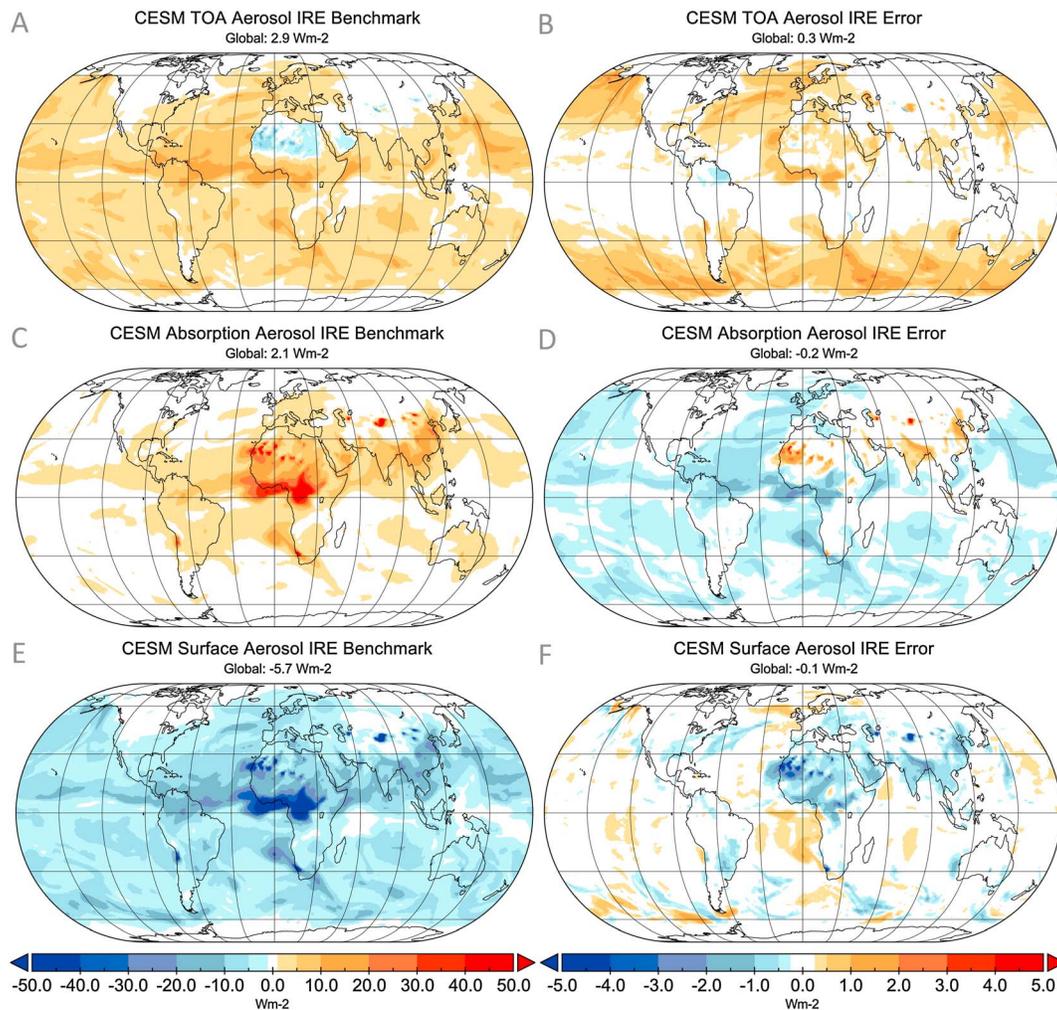
Aerosol IRE and its error have been quantified for interim versions of two climate models that are planning to participate in CMIP6: AM4 and CESM 1.2.2. Since neither aerosol properties nor their distribution has been prescribed, a unique benchmark is created for each model. The errors characterized here arise only from the radiative parameterization’s handling of radiative transfer and not from errors in aerosol distribution, emission, concentration, or conversion from physical properties to optical properties or the representation of surface albedo. Visualization of these benchmark diagnostics are presented both globally and as a



**Figure 3.** Diagnostics of aerosol IRE and error for GFDL AM4. (a) Diurnally averaged shortwave aerosol IRE for net upward flux at the TOA for vernal equinox, 2005. (b) Error in Figure 3a defined as AM4 minus benchmark calculations. (c) Same as Figure 3a but for atmospheric absorption. (d) Same as Figure 3b but for atmospheric absorption. (e) Same as Figure 3a but for downward flux at surface. (f) Same as Figure 3b but for downward flux at surface.

function of aerosol optical properties, neither of which were available in previous studies, allowing for hypothesis formulation of error sources.

The global nature of these benchmarks allows for a comprehensive error diagnostic calculation, where the average error in TOA, surface, and atmospheric absorption aerosol IRE is a true global daily average, not just the average of a few columns. This paradigm stands in contrast to the last two major intercomparisons of aerosol IRE errors among GCM radiative parameterizations. CIRC (Oreopoulos et al., 2012) had four single-column cloud-free cases with and without aerosols, while AeroCOM (Myhre et al., 2012) compared either a scattering aerosol or an absorbing aerosol at two zenith angles also giving a total of four single-column cases. To demonstrate the problems associated with using a limited number of cases to diagnose errors in aerosol IRE, we randomly select four cases from a single time step of the CESM model: two with low zenith angle (one low aerosol and one medium aerosol case) and two with high zenith angle (one medium and one high aerosol case). In Figure 2 the average error from these four selected profiles and globally and diurnally averaged error are included for CESM 1.2.2 and AM4. At TOA for AM4 and surface for CESM even the sign of the global, daily average error is not captured by the sample average. In addition to better sampling the diversity of meteorological and aerosol conditions, the spatial pattern of error can be explored to provide context and variability for the global average error. These errors are illustrated globally in Figures 3 and 4 and reveal the extent of the information that is lost by using only a small number of cases.

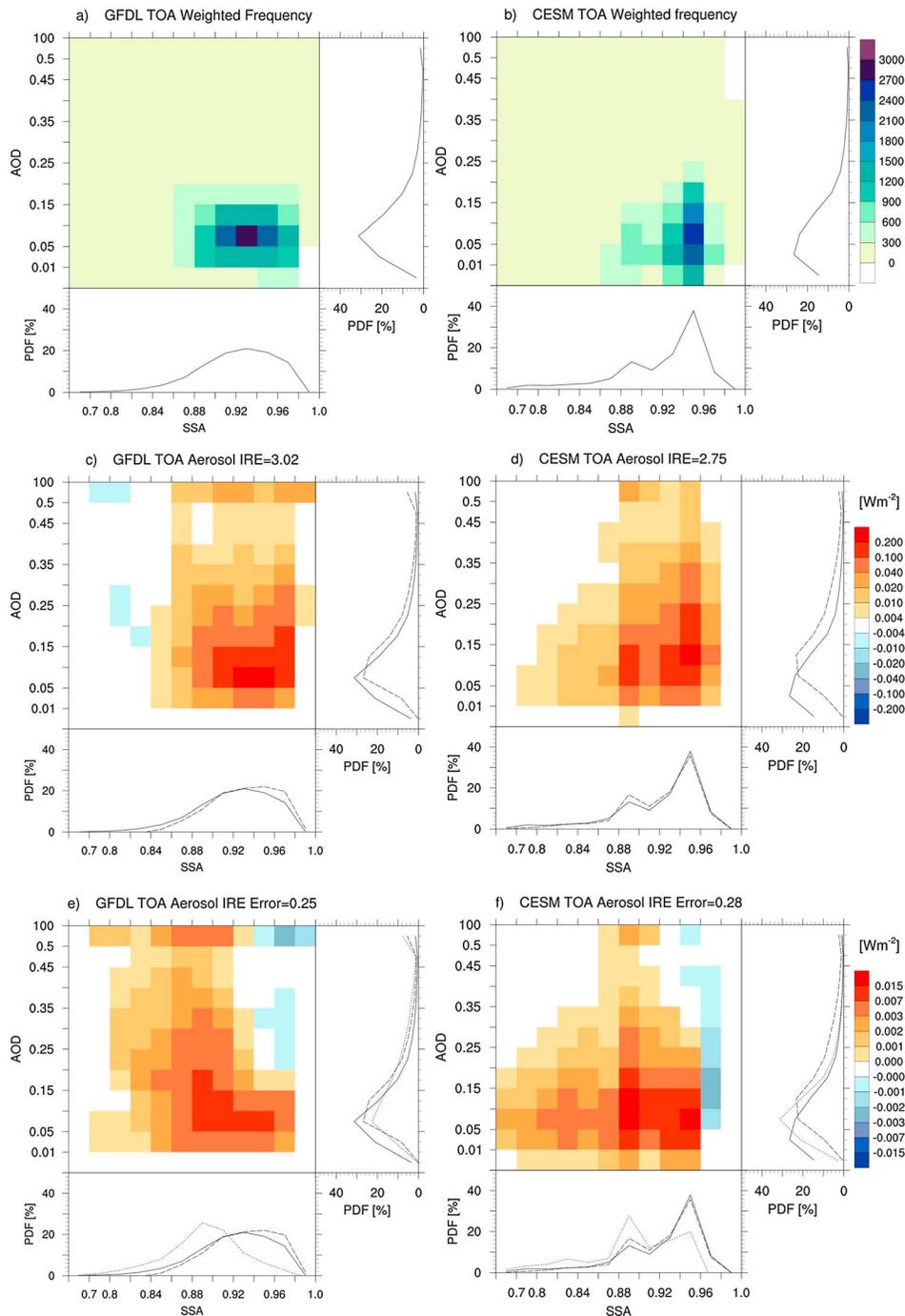


**Figure 4.** Diagnostics of aerosol IRE and error for CESM 1.2.2. (a) Diurnally averaged shortwave aerosol IRE for net upward flux at the TOA for vernal equinox, 2005. (b) Error in Figure 4a defined as CESM minus benchmark calculations. (c) Same as Figure 4a but for atmospheric absorption. (d) Same as Figure 4b but for atmospheric absorption. (e) Same as Figure 4a but for downward flux at the surface. (f) Same as Figure 4b but for downward flux at surface.

### 3.1. Spatial Distribution of IRE Error

The spatial patterns of aerosol IRE and associated error for the AM4 and CESM benchmark calculations for net upward flux at the top of the atmosphere, atmospheric absorption, and downward flux at the surface are shown in Figures 3 and 4. They reveal the spatial heterogeneity of aerosol IRE, arising from the variety of meteorological and aerosol conditions. For example, over the Sahara and Arabian Desert regions and nearby waters, the presence of aerosol reduces the reflection back to space over the reflective land surfaces but leads to increased reflection back to space over the less reflective Mediterranean Sea and Atlantic Ocean (Figures 3a and 4a).

Both AM4 and CESM's radiative parameterizations overestimate the global mean radiation reflected to space by their own aerosols (Figures 3a and 4b) by  $0.2 \text{ W/m}^2$  and  $0.3 \text{ W/m}^2$ , respectively, and underestimate their aerosol's global mean absorption by  $0.9 \text{ W/m}^2$  and  $0.2 \text{ W/m}^2$ , respectively (Figures 3d and 4d). In the regions of AM4's strongest aerosol absorption (Figure 3c) IRE errors exceed  $5 \text{ W/m}^2$  (Figure 3d) and the global mean represents an error of 20%. AM4's aerosol IRE errors at the TOA and for absorption in the atmosphere (Figures 3b and 3d) are both highly correlated with the aerosol IRE on atmospheric absorption (Figure 3c) with  $R^2 = 0.908$  and  $R^2 = 0.791$ , respectively. However, the same is not true for CESM, which has much lower correlations of atmospheric absorption aerosol IRE with the TOA IRE error ( $R^2 = 0.029$ ) and absorption IRE error ( $R^2 = 0.0002$ ). This highlights the need for individual benchmark calculations to get at the differing relationships between IRE, error, and aerosol optical properties for each model.



**Figure 5.** Benchmark diagnostics of (a and b) aerosol frequency distributions, weighted by net upward TOA cloud-free, aerosol-free flux, (c and d) aerosol IRE for net upward flux at the TOA, and (e and f) aerosol IRE error at the TOA, defined as GCM minus benchmark calculations as functions of column aerosol single-scattering albedo (SSA) and optical depth (AOD) across all bands and time steps for GFDL AM4 (Figures 5a, 5c, and 5e) and CESM (Figures 5b, 5d, and 5f). The corresponding univariate PDFs of frequency (solid), aerosol IRE (dashed), and aerosol IRE error (dotted) are shown on the bottom (as a function of SSA) and right (as a function of AOD) axes of each plot.

Aerosol IRE at the surface is overestimated by AM4 and slightly underestimated by CESM (Figures 3f and 4f) with global means of  $1.0 \text{ W/m}^2$  and  $-0.1 \text{ W/m}^2$ , respectively. CESM's small global average error is partly the result of compensating regional effects and partly due to a smaller root-mean-square (RMS) error: CESM's RMS error is  $0.4 \text{ W/m}^2$  compared to AM4's  $1.3 \text{ W/m}^2$ . The fact that AM4's error in absorption (Figure 3d) and error at the surface (Figure 3f) are much greater than at TOA (Figure 3b) indicates that the

predominant problem with the AM4 radiative parameterization is excessive transmission of incoming solar radiation through aerosols to the surface that should be absorbed in the atmospheric column, rather than excessive reflection by aerosols back to space.

Together, Figures 2–4 show that the use of a small number of profiles in either of these models is potentially biased relative to true global, diurnal averages. A simple Monte Carlo analysis indicates that over 1,500 profiles are needed to estimate global IRE values to within  $0.1 \text{ W/m}^2$ , and while more advanced techniques are possible (Kroese et al., 2011), they require model-specific optimization. Furthermore, even with an unbiased global average error, there is no dispositive information or at least information that can be used to test hypotheses about sources of model IRE error.

### 3.2. Error as a Function of Aerosol Optical Properties

From the same diagnostics, we can also evaluate IRE errors as a function of model-specific aerosol optical properties. This not only reveals whether parameterizations have problems with certain aerosol characteristics but also summarizes and compares the distributions of aerosol optical properties. Figure 5 shows an analysis of aerosol IRE and error at TOA as a function of column aerosol optical properties, namely, AOD and SSA. Similar plots for atmospheric absorption and at the surface can be found in the supporting information. The frequency is weighted by the cloud-free, aerosol-free flux at the TOA to preference aerosols that may have the largest influence on TOA flux (Figures 5a and 5b). We also explicitly account for the spectrally varying contributions to aerosol IRE and its associated error by representing each spectral band's optical properties, aerosol IRE, and associated error on the same plot (Figures 5c–5f), rather than ascribing broadband IRE and IRE error to visible-band optical properties (as is often the case in previous studies). Each bin represents the portion of the spectrally integrated global and diurnal average contributed by the model columns with optical properties falling within the bin's bounds. Therefore, summing each bin's values results in the spectrally integrated, global, diurnal average.

Our results show that for both models, the bulk of the globe is covered by aerosols with optical depth less than 0.1 (Figures 5a and 5b). Single-scattering albedo peaks at 0.92–0.94 for AM4 but is bimodal for CESM with a maximum peak at 0.94–0.96. For CESM the peak in the aerosol IRE probability density function (PDF) is shifted to higher AOD (Figure 5d dashed line), relative to the AOD frequency distribution (solid line), whereas the contribution to IRE for AM4 is proportional to the distribution of optical properties (Figure 5c). For both AM4 and CESM the distribution of aerosol IRE error is shifted toward lower SSA, relative to the PDFs of IRE and aerosol SSA frequency (Figures 5e and 5f). These panels also indicate that varying degrees of error compensation exists between aerosols of differing SSA: the negative error values (in blue) for AM4 and especially for CESM show that high SSA aerosols' contribution to the IRE error is of opposite sign from the average effect. These results indicate that the contribution to the error by aerosols of differing AOD and SSA differs between the two models, as shown in Figures 5d and 5f. Experiments using prescribed aerosols to uncover errors would need to project their results onto the model's actual distribution of optical properties in order to characterize the relevance of their estimations of errors to a participating model's true error. No such projection is required with the diagnostics presented here.

## 4. Conclusions

The persistent range in direct natural and anthropogenic aerosol forcing across successive phases of CMIP demands a novel approach to diagnosing model error sources. The method presented here provides new and targeted diagnostics for aerosol IRE in a way that leverages the rising computational infrastructure that can now solve line-by-line radiative transfer with scattering on a model's native spatial grid. These resulting calculations can be used to infer information about and ultimately test hypotheses on, sources of radiative parameterization IRE error. Additionally, the method can be readily adapted to evaluate the contribution of known error sources including the coarse spectral and angular representation of surface reflectance and condensate optics to aerosol IRE error.

We have presented here proof-of-concept calculations and diagnostics for RFMIP Aerosol-IRF, which already reveal some disquieting results for AM4 and CESM. Using a unique protocol, we tested these models' radiation schemes over a wide range of meteorological, surface, and aerosol conditions and revealed subtle errors in IRE that are not readily apparent in the global mean or a small subset of columns. At TOA the global

average IRE errors due only to radiative parameterizations are the same order of magnitude as the anthropogenic direct aerosol forcing reported in AR4 and AR5 (Myhre et al., 2013). Since aerosol absorption has a direct impact upon the hydrological cycle (Ming et al., 2010; Persad et al., 2012; Pendergrass & Hartmann, 2012), the fact that AM4 is missing 20% of its aerosol absorption implies that its hydrological response to aerosol forcing is likewise too weak. For CESM, compensating errors suggest that the model's change in evaporation due to aerosol forcing will be too weak over the ocean but too strong over land. These errors must be addressed within each model to more accurately simulate the climate response to aerosols.

This targeted diagnostic approach is distinct from, and complementary to, other aerosol model and radiative parameterization intercomparison exercises, because it evaluates models without prescribing aerosol emissions or distributions. Working with the diagnostics demonstrated here, modeling centers can evaluate the sources of error and their importance to the model. By participating in RFMIP Aerosol-IRF, we expect that the modeling centers will receive detailed error information that they can then use to target development efforts to reduce their aerosol forcing error. Modeling centers interested in participating should review the steps to participation listed on the RFMIP-Aerosol-IRF webpage (<https://goo.gl/BhUVdj>).

An additional benefit is that with broad participation, we can quantify the true range of aerosol forcing for CMIP6 without the artifacts of radiative parameterization errors. Given the disparate results of these two models when running with prescribed aerosol emissions, this benefit will be timely for CMIP6. Ultimately, novel approaches for multimodel diagnostics such as the one discussed here need to be considered so that model errors are identified and resolved. New methods are needed to reduce uncertainty in climate model projections, and the diagnostics described here support an approach that ensures that the range of reported model results is due to climate feedback processes (structural and parametric uncertainty) rather than to forcing error.

#### Acknowledgments

A. L. Jones is funded by DOE award DE-SC0012313, "Collaborative Research: Centralized activities in support of the Radiative Forcing Model Intercomparison Project" and Princeton University's Cooperative Institute for Climate Studies (CICS). Computational efforts were completed on Gaea at Oak Ridge National Laboratory. D. R. Feldman was supported by the Director, Office of Science and the Regional and Global Climate Modeling Program of the U.S. Department of Energy and used resources of the National Energy Research Scientific Computing Center (NERSC), also supported by the Office of Science of the U.S. Department of Energy, under contract DE-AC02-05CH11231. R. Pincus is supported by Regional and Global Climate Modeling Program of the U.S. Department of Energy under grant DE-SC0012549. Data is permanently archived at GFDL and can be obtained upon request. A. L. Jones and D. R. Feldman are co-first authors and have contributed equally.

#### References

- Andreae, M., Jones, C. D., & Cox, P. M. (2005). Strong present-day aerosol cooling implies a hot future. *Nature*, *435*(7046), 1187–1190. <https://doi.org/10.1038/nature03671>
- Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., ... Zhang, X. Y. (2013). Clouds and aerosols. In T. F. Stocker, et al. (Eds.), *Climate change 2013: The physical science basis. Contribution of working group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 571–657). Cambridge, United Kingdom and New York: Cambridge University Press.
- Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., ... Brown, P. D. (2005). Atmospheric radiative transfer modeling: A summary of the AER codes. *Journal of Quantitative Spectroscopy and Radiation Transfer*, *91*(2), 233–244. <https://doi.org/10.1016/j.jqsrt.2004.05.058>
- Collins, W. J., Lamarque, J. F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., ... Smith, S. J. (2017). AerChemMIP: Quantifying the effects of chemistry and aerosols in CMIP6. *Geoscientific Model Development*, *10*(2), 585–607. <https://doi.org/10.5194/gmd-10-585-2017>
- Donner, L. J., Wyman, B. L., Hemler, R. S., Horowitz, L. W., Ming, Y., Zhao, M., ... Zeng, F. (2011). The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *Journal of Climate*, *24*(13), 3484–3519. <https://doi.org/10.1175/2011JCLI3955.1>
- Dudhia, A. (2017). The reference forward model. *Journal of Quantitative Spectroscopy and Radiation Transfer*, *186*, 243–253. <https://doi.org/10.1016/j.jqsrt.2016.06.018>
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>
- Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D. W., ... Van Dorland, R. (2007). Changes in atmospheric constituents and in radiative forcing. In S. Solomon, et al. (Eds.), *Climate change 2007: The physical science basis. Contribution of working group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 129–234). Cambridge, United Kingdom and New York: Cambridge University Press.
- Freidenreich, S., & Ramaswamy, V. (2005). Refinement of the Geophysical Fluid Dynamics Laboratory solar benchmark computations and an improved parameterization for climate models. *Journal of Geophysical Research*, *110*, D17105. <https://doi.org/10.1029/2004JD005471>
- Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., ... Marshall, S. (2013). The Community Earth System Model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, *94*(9), 1339–1360. <https://doi.org/10.1175/BAMS-D-12-00121.1>
- Kinne, S., O'Donnel, D., Stier, P., Kloster, S., Zhang, K., Schmidt, H., ... Stevens, B. (2013). MAC-v1: A new global aerosol climatology for climate studies. *Journal of Advances in Modeling Earth Systems*, *5*(4), 704–740. <https://doi.org/10.1002/jame.20035>
- Kroese, D. P., Taimre, T., & Botev, Z. I. (2011). *Handbook of Monte Carlo methods*. Hoboken, NJ: John Wiley & Sons. <https://doi.org/10.1002/9781118014967>
- Mann, G. W., Carslaw, K. S., Reddington, C. L., Pringle, K. J., Schulz, M., Asmi, A., ... Henzing, J. S. (2014). Intercomparison and evaluation of global aerosol microphysical properties among AeroCom models of a range of complexity. *Atmospheric Chemistry and Physics*, *14*(9), 4679–4713. <https://doi.org/10.5194/acp-14-4679-2014>
- Meehl, G. A., Covey, C., Taylor, K. E., Delworth, T., Stouffer, R. J., Latif, M., ... Mitchell, J. F. B. (2007). The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bulletin of the American Meteorological Society*, *88*(9), 1383–1394. <https://doi.org/10.1175/BAMS-88-9-1383>
- Ming, Y., Ramaswamy, V., & Persad, G. (2010). Two opposing effects of absorbing aerosols on global-mean precipitation. *Geophysical Research Letters*, *37*, L13701. <https://doi.org/10.1029/2010GL042895>

- Moncet, J. L., & Clough, S. A. (1997). Accelerated monochromatic radiative transfer for scattering atmospheres: Application of a new model to spectral radiance observations. *Journal of Geophysical Research*, 102(D18), 21,853–21,866. <https://doi.org/10.1029/97JD01551>
- Myhre, G., Samset, B. H., Schulz, M., Balkanski, Y., Bauer, S., Bernsten, T. K., ... Bian, H. (2012). Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations. *Atmospheric Chemistry and Physics*, 12, 22,355–22,413. <https://doi.org/10.5194/acp-13-1853-2013>
- Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestad, J., Huang, J., ... Koch, D. (2013). Anthropogenic and natural radiative forcing. In D. Jacob, A. J. Ravishankara, & K. Shine (Eds.), *Climate change 2013: The physical science basis. Contribution of working group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 659–740). Cambridge, United Kingdom and New York: Cambridge University Press.
- Oreopoulos, L., Mlawer, E., Delamere, J., Shippert, T., Cole, J., Fomin, B., ... Rossow, W. B. (2012). The continual intercomparison of radiation codes: Results from phase I. *Journal of Geophysical Research*, 117, D06118. <https://doi.org/10.1029/2011JD016821>
- Pendergrass, A., & Hartmann, D. L. (2012). Global-mean precipitation and black carbon in AR4 simulations. *Geophysical Research Letters*, 39, L01703. <https://doi.org/10.1029/2011GL050067>
- Persad, G., Ming, Y., & Ramaswamy, V. (2012). Tropical tropospheric-only responses to absorbing aerosols. *Journal of Climate*, 25(7), 2471–2480. <https://doi.org/10.1175/JCLI-D-11-00122.1>
- Pincus, R., Forster, P. M., & Stevens, B. (2016). The Radiative Forcing Model Intercomparison Project (RFMIP): Experimental protocol for CMIP6. *Geoscientific Model Development*, 9(9), 3447–3460. <https://doi.org/10.5194/gmd-9-3447-2016>
- Ramanathan, V., & Carmichael, G. (2008). Global and regional climate changes due to black carbon. *Nature Geoscience*, 1(4), 221–227. <https://doi.org/10.1038/ngeo156>
- Ramaswamy, V., & Freidenreich, S. (1991). Solar radiative line-by-line determination of water vapor absorption and water cloud extinction in inhomogeneous atmospheres. *Journal of Geophysical Research: Atmospheres*, 96(D5), 9133–9157. <https://doi.org/10.1029/90JD00083>
- Randles, C. A., Kinne, S., Myhre, G., Schulz, M., Stier, P., Fischer, J., ... Lu, P. (2013). Intercomparison of shortwave radiative transfer schemes in global aerosol modeling: Results from the AeroCom Radiative Transfer Experiment. *Atmospheric Chemistry and Physics*, 13(5), 2347–2379. <https://doi.org/10.5194/acp-13-2347-2013>
- Rothman, L. S., Gordon, I. E., Babikov, Y., Barbe, A., Chris Benner, D., Bernath, P. F., ... Birk, M. (2013). The HITRAN2012 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiation Transfer*, 130, 4–50. <https://doi.org/10.1016/j.jqsrt.2013.07.002>
- Santer, B. D., Bonfils, C., Painter, J. F., Zelinka, M. D., Mears, C., Solomon, S., ... Wentz, F. J. (2014). Volcanic contribution to decadal changes in tropospheric temperature. *Nature Geoscience*, 7(3), 185–189. <https://doi.org/10.1038/ngeo2098>
- Shindell, D. T., Lamarque, J. F., Schulz, M., Flanner, M., Jiao, C., Chin, M., ... Lo, F. (2013). Radiative forcing in the ACCMIP historical and future climate simulations. *Atmospheric Chemistry and Physics*, 13(6), 2939–2974. <https://doi.org/10.5194/acp-13-2939-2013>
- Solomon, S., Daniel, J. S., Neely, R. R., Vernier, J. P., Dutton, E. G., & Thomason, L. W. (2011). The persistently variable “background” stratospheric aerosol layer and global climate change. *Science*, 333(6044), 866–870. <https://doi.org/10.1126/science.1206027>
- Stevens, B. (2015). Rethinking the lower bound on aerosol radiative forcing. *Journal of Climate*, 28(12), 4794–4819. <https://doi.org/10.1175/JCLI-D-14-00656.1>
- Stevens, B., Fiedler, S., Kinne, S., Peters, K., Rast, S., Müsse, J., ... Mauritsen, T. (2017). MACv2-SP: A parameterization of anthropogenic aerosol optical properties and associated Twomey effect for use in CMIP6. *Geoscientific Model Development*, 10(1), 433–452. <https://doi.org/10.5194/gmd-10-433-2017>
- Stier, P., Schutgens, N. A. J., Bellouin, N., Bian, H., Boucher, O., Chin, M., ... Zhou, C. (2013). Host model uncertainties in aerosol radiative forcing estimates: Results from the AeroCom prescribed intercomparison study. *Atmospheric Chemistry and Physics*, 13(6), 3245–3270. <https://doi.org/10.5194/acp-13-3245-2013>
- Stouffer, R. J., Eyring, V., Meehl, G. A., Bony, S., Senior, C., Stevens, B., & Taylor, K. E. (2017). CMIP5 scientific gaps and recommendations for CMIP6. *Bulletin of the American Meteorological Society*, 98(1), 95–105. <https://doi.org/10.1175/BAMS-D-15-00013.1>
- Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4), 485–498. <https://doi.org/10.1175/BAMS-D-11-00094.1>
- Zhao, M., Golaz, J. C., Held, I. M., Ramaswamy, V., Lin, S. J., Ming, Y., ... Guo, H. (2016). Uncertainty in model climate sensitivity traced to representations of cumulus precipitation microphysics. *Journal of Climate*, 29(2), 543–560. <https://doi.org/10.1175/JCLI-D-15-0191.1>